

## Methods

# Estimation of time-varying causal effects with multivariable Mendelian randomization: some cautionary notes

Haodong Tian <sup>1\*</sup> and Stephen Burgess <sup>1,2</sup>

<sup>1</sup>MRC Biostatistics Unit, School of Clinical Medicine, University of Cambridge, Cambridge, UK and

<sup>2</sup>British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

\*Corresponding author. MRC Biostatistics Unit, University of Cambridge, East Forvie Building, Forvie Site, Robinson Way, Cambridge, CB2 0SR, UK. E-mail: [haodong.tian@mrc-bsu.cam.ac.uk](mailto:haodong.tian@mrc-bsu.cam.ac.uk)

Received 4 March 2022; Editorial decision 8 December 2022; Accepted 21 December 2022

## Abstract

**Introduction:** For many exposures present across the life course, the effect of the exposure may vary over time. Multivariable Mendelian randomization (MVMR) is an approach that can assess the effects of related risk factors using genetic variants as instrumental variables. Recently, MVMR has been used to estimate the effects of an exposure during distinct time periods.

**Methods:** We investigated the behaviour of estimates from MVMR in a simulation study for different time-varying causal scenarios. We also performed an applied analysis to consider how MVMR estimates of body mass index on systolic blood pressure vary depending on the time periods considered.

**Results:** Estimates from MVMR in the simulation study were close to the true values when the outcome model was correctly specified: i.e. when the outcome was a discrete function of the exposure at the precise time points at which the exposure was measured. However, in more realistic cases, MVMR estimates were misleading. For example, in one scenario, MVMR estimates for early life were clearly negative despite the true causal effect being constant and positive. In the applied example, estimates were highly variable depending on the time period in which genetic associations with the exposure were estimated.

**Conclusions:** The poor performance of MVMR to study time-varying causal effects can be attributed to model misspecification and violation of the exclusion restriction assumption. We would urge caution about quantitative conclusions from such analyses and even qualitative interpretations about the direction, or presence or absence, of a causal effect during a given time period.

**Key words:** Instrumental variables, causal inference, life-course epidemiology, misspecification, exclusion restriction

**Key Messages**

- Multivariable Mendelian randomization (MVMR) is a popular analysis approach that has recently been applied to analyse the effect of a single exposure during different time periods, in order to understand the time-varying effect of the exposure.
- Our simulation study showed that estimates from MVMR can be misleading when the model relating the exposure to the outcome is misspecified.
- Misspecification is likely to be common in practice, as it is implausible that the outcome is a discrete function of the values of the exposure during the specific time periods when the exposure was measured.
- MVMR estimates can be unreliable as a guide to either the direction, or even the existence, of a causal effect of the exposure during a particular time period.
- These cautions are particularly relevant to the case that the risk factors in the analysis are not distinct exposures, but measurements of the same exposure at different time points.

**Introduction**

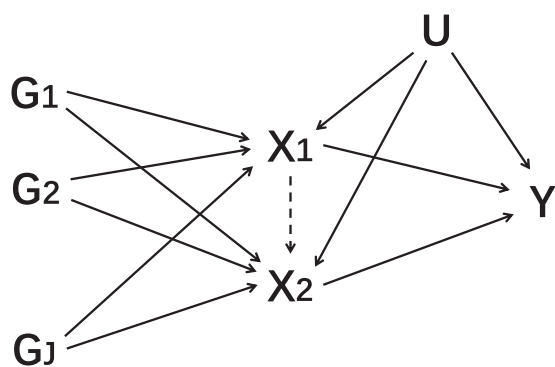
Multivariable Mendelian randomization (MVMR) is an extension of standard (univariable) Mendelian randomization to investigate the causal effects of related risk factors with shared genetic predictors.<sup>1</sup> In standard Mendelian randomization, we take genetic variants that are predictors of a single risk factor, and assess whether genetically predicted levels of the risk factor are associated with the outcome.<sup>2</sup> If the genetic variants satisfy the instrumental variable assumptions, then an association between genetically predicted levels of the risk factor and the outcome is indicative of a causal effect of the risk factor on the outcome.<sup>3</sup> MVMR is analogous, except that instead of testing whether genetically predicted values of a single risk factor are associated with the outcome, we test whether genetically predicted values of multiple risk factors are associated with the outcome or not in a multivariable model.<sup>4</sup> Estimates from MVMR represent direct effects of the risk factors on the outcome.<sup>5</sup> The method is primarily used in two contexts: first, to assess the effect of an exposure when genetic variants associated with the exposure may have pleiotropic effects on the outcome via other measured risk factors;<sup>1</sup> and second, to assess the relative contribution of causal pathways from the exposure to the outcome via other measured risk factors in a mediation analysis.<sup>5,6</sup>

We investigate a related application of MVMR that has been considered in empirical investigations<sup>7,8</sup> and has recently been described from a methodological perspective<sup>9,10</sup> that we refer to as time-varying Mendelian randomization. In this setting, the risk factors are not separate exposures, but rather multiple measures of the same exposure at different time points (Figure 1). For instance, Richardson *et al.* considered body mass index (BMI) measured during early life and during later life as separate risk factors, and assessed whether genetically predicted values

of early-life and later-life BMI were associated with coronary artery disease (CAD) risk.<sup>7</sup> The authors interpreted a positive univariable association between genetically predicted early-life BMI and CAD risk as evidence that early-life BMI is a causal risk factor for CAD, and lack of an independent association between genetically predicted early-life BMI and CAD risk in a multivariable model that also included genetically predicted later-life BMI as evidence that early-life BMI does not have a direct effect on CAD risk, but that the risk is mediated via later-life BMI.

These analyses pose several difficulties. First, it is necessary to have some linear independence between genetic predictors of the exposure at different time points.<sup>11</sup> If genetic associations with the exposure at different time points are perfectly proportional, then it is not possible to disentangle the effects of the exposure at different time points. It is not necessary to find distinct genetic predictors of the exposure at different time points, but if correlations across variants between the genetic associations with the exposure at different time points are strong, then the analysis will have limited utility.

However, even when such genetic variants are available, the use of MVMR to investigate time-varying causal effects necessitates strong parametric assumptions that are unlikely to be plausible in practice. In univariable Mendelian randomization, in order to estimate a parameter that represents a causal effect, it is necessary to make parametric assumptions.<sup>12</sup> There are different parametric assumptions, depending on which causal effect to be identified or interpreted. For example, for the average treatment effect (ATE), one normally requires the homogeneity and the linearity of exposure–outcome model.<sup>13</sup> Similarly, monotonicity is usually required for the local average treatment effect (LATE). However, even if these assumptions are not satisfied, the standard Mendelian randomization



**Figure 1** Directed acyclic graph illustrating multivariable Mendelian randomization assumptions for two risk factors ( $X_1$  and  $X_2$ ),  $J$  genetic variants ( $G_1, G_2, \dots, G_J$ ) that are assumed to satisfy the multivariable instrumental variable assumptions, an outcome ( $Y$ ) and an unmeasured confounder ( $U$ ). In time-varying Mendelian randomization, we assume that the two risk factors are measurements of the same exposure at different time points. The dashed arrow from  $X_1$  to  $X_2$  indicates the potential causal effect of the exposure at time point 1 on the exposure at time point 2

estimate is a valid test statistic for the sharp causal null hypothesis that the exposure has no causal effect on the outcome at any time point.<sup>14</sup> Rejection of this null hypothesis implies that the exposure has a causal effect on the outcome at some time point in at least a subset of the population. Here, we will show by counterexample that an analogous result for time-varying Mendelian randomization does not hold.

In this paper, we will explore the behaviour of estimates from MVMR for a time-varying exposure in a simple simulation study. We shall show that, when the model for the outcome is correctly specified, estimates from MVMR are close to the true parameter values. However, when the model is incorrectly specified (as is likely in practice), estimates are highly misleading. We also perform an applied MVMR analysis of BMI on systolic blood pressure (SBP) and show that estimates and inferences are highly variable depending on the time period at which genetic associations with BMI are estimated.

## Methods

MVMR has been described at length in the literature<sup>1,4</sup> and in particular in the context of time-varying causal effects.<sup>9,10</sup> We here provide a brief overview of the approach.

### Assumptions and estimation

We use the term ‘exposure’ to describe the putative causal factor and the term ‘risk factor’ to describe the value of the exposure at a particular time point. MVMR requires each

genetic variant to satisfy the multivariable instrumental variable assumptions<sup>15</sup>:

- the variant is associated with one or more of the risk factors,
- the variant is not associated with the outcome via a confounding pathway and
- the variant does not affect the outcome directly, only possibly indirectly via one or more of the risk factors.

In order to estimate causal effects, we need to make additional assumptions. For simplicity, we assume that all variables are continuous, and the associations between the genetic variants and the risk factors, the genetic variants and the outcome, and the causal effect of the risk factors on the outcome are homogeneous and linear without effect modification by any confounder.<sup>3</sup> This enables the estimation of the direct effects of the risk factors on the outcome using the two-stage least squares (2SLS) method, which can be implemented by first regressing the risk factors on the genetic variants, and then regressing the outcome on the fitted values of the risk factors in a multivariable model.<sup>16</sup>

The term ‘direct effect’ is imprecise; it has previously been argued that the direct effects estimated in instrumental variable analyses are most naturally interpreted as controlled direct effects (i.e. the effect of varying the exposure while fixing the mediator at a given value), as the instrumental variables set the values of the risk factors.<sup>6</sup> However, in the all-linear setting, natural and controlled direct effects take the same value, so any difference is a question of interpretation. An alternative estimation method that could be applied here is g-estimation of structural mean models, although estimates are similar to those from 2SLS in the all-linear setting.<sup>10</sup> Equivalent estimates to those from 2SLS would also be obtained from the multivariable inverse-variance weighted method if we had access to summarized data on the genetic associations with the outcome and with the exposure at the relevant time points.<sup>17</sup>

### Simulation studies

We investigate the behaviour of MVMR estimates for a time-varying exposure in a series of simulation studies. We simulate the time-varying exposure  $X(t)$  according to the following data-generating model:

$$X(t) = \sum_{j=1}^{30} \alpha_j(t) G_j + \cos(t)U + \sin(t)\epsilon_X \quad (1)$$

where the genetic variants  $G_j$  for  $j = 1, 2, \dots, 30$  follow independent binomial distributions  $B(2, 0.3)$ , and the confounder  $U$  and the error term  $\epsilon_X$  have independent standard normal distributions. The exposure varies over

time  $t$ , which can be interpreted as an individual's age, owing to sinusoidal effects of the error term  $[\sin(t)]$ , the confounder  $[\cos(t)]$  and the genetic variants:

$$\alpha_j(t) = A_{1,j} + A_{2,j} \cos(A_{3,j}t - A_{4,j}) \quad (2)$$

where  $A_{1,j}$ ,  $A_{2,j}$ ,  $A_{3,j}$  and  $A_{4,j}$  all have independent normal distributions for each variant  $j$ . We use the cosine function to generate variability in the genetic effects and thus avoid weak instruments bias caused by the high linear dependence between genetic predictors of the exposure at different time points. The distributions of  $\{A_{1,j}, A_{2,j}, A_{3,j}, A_{4,j}\}$  are fixed for each scenario, but vary between scenarios. Detailed values for the independent normal distribution are shown in [Supplementary Table S1](#) (available as [Supplementary data](#) at *IJE* online).

We consider two scenarios for the outcome. First, we assume that the outcome is a function of the exposure at two fixed time points,  $t = 10$  and  $t = 50$ :

$$Y = 0.4X(10) - 0.8X(50) + U + \epsilon_Y \quad (3)$$

where  $U$  is the confounder as above and  $\epsilon_Y$  has an independent standard normal distribution. The true causal effect of the exposure at time 10 is +0.4 and the causal effect of the exposure at time 50 is -0.8. We consider estimates from MVMR in four cases: first, when the exposure is measured at times 10 and 50 (Scenario 1A); then when the exposure is measured at times 10, 40 and 50 (Scenario 1B); at times 15 and 30 (Scenario 1C); and at times 15 and 50 (Scenario 1D). We recognize that this outcome model is somewhat unrealistic; we explore these scenarios to investigate whether methods can consistently estimate causal effects when the model is correctly specified.

Second, and more realistically, we assume that the outcome is a continuous function of the exposure that varies over time. We express the relationship between the

exposure and outcome using an integral, where the causal effect of the exposure depends on the time-varying function  $\beta(t)$ :

$$Y = \int_0^{50} \beta(t)X(t)dt + U + \epsilon_Y \quad (4)$$

where again the confounder  $U$  and  $\epsilon_Y$  have independent standard normal distributions. We consider three different scenarios for  $\beta(t)$ :

$$\beta(t) = \begin{cases} 0 & \text{for } t \in [0, 40] \\ 1 & \text{for } t \in [40, 50] \end{cases} \text{ (Scenario 2A)} \quad (5a)$$

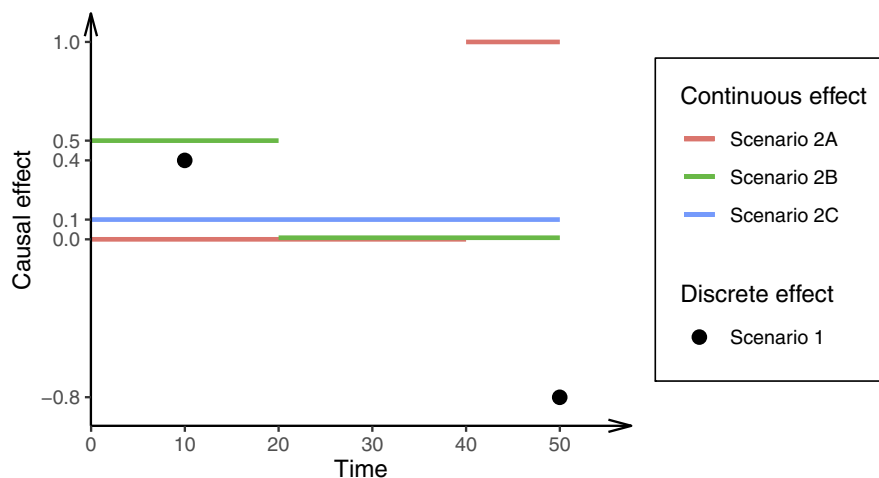
$$\beta(t) = \begin{cases} 0.5 & \text{for } t \in [0, 20] \\ 0 & \text{for } t \in [20, 50] \end{cases} \text{ (Scenario 2B)} \quad (5b)$$

$$\beta(t) = 0.1 \quad \text{for } t \in [0, 50] \text{ (Scenario 2C).} \quad (5c)$$

In Scenario 2A, the causal effect of the exposure is null in early life (up to time 40) and positive in later life. In Scenario 2B, the causal effect of the exposure is positive in early life (up to time 20) and null in later life. In Scenario 2C, the causal effect of the exposure is constant and positive across the life course. The model parameters are illustrated in [Figure 2](#). We consider MVMR estimates when the risk factors are the exposure measured at times 10 and 50. In Scenario 2C, we also consider a wider range of choices of timings for the exposure measurements and consider the impact on estimates.

### Interpreting multivariable estimates as causal effects

In order to better interpret parameters estimated in time-varying MVMR, we consider a simplified scenario in which the genetic effects on the exposure vary linearly over time. Suppose we have two measured time points  $t_1$  and  $t_2$



**Figure 2** The model parameters for each scenario in the simulation. For continuous effects, the outcome is a continuous function of the exposure that varies over time (0, 50). For the discrete effect, the outcome is a function of the exposure at the two fixed time points (10 and 50)

satisfying  $0 \leq t_1 < t_2 \leq 50$ , and consider estimates at the measured time points  $\hat{\beta}_1$  and  $\hat{\beta}_2$  for an outcome  $Y$  defined by Equation (4). The total cumulative effect  $\int_0^{50} \beta(t) dt$  represents the impact of a lifelong increase in the exposure by one unit. These estimates have the asymptotic values (see [Supplementary Text S1](#) and [S2](#), available as [Supplementary data](#) at *IJE* online):

$$\begin{aligned} \text{plim}_{n \rightarrow \infty} \hat{\beta}_1 &= \int_0^{50} \beta(t) \frac{t_2 - t}{t_2 - t_1} dt \quad \text{and} \\ \text{plim}_{n \rightarrow \infty} \hat{\beta}_2 &= \int_0^{50} \beta(t) \frac{t - t_1}{t_2 - t_1} dt \end{aligned} \quad (6)$$

which represent weighted cumulative effects across the whole life course. Hence the estimates do not represent an effect limited to a particular time period: the estimate at  $t_1$  more strongly represents the effect of the exposure during early life and the estimate at  $t_2$  more strongly represents the effect during later life, but both estimates are influenced by the effect of the exposure across the whole life course. Hence, if there is a sustained effect of the exposure during early life, the estimate at  $t_2$  will be non-zero even if the effect during later life is null.

If  $t_1 = 0$  and  $t_2 = 50$ , then the weighting functions  $\frac{t_2 - t}{t_2 - t_1}$  and  $\frac{t - t_1}{t_2 - t_1}$  take values between 0 and 1, and the estimates at  $t_1$  and  $t_2$  will be non-negative if the effect function  $\beta(t)$  is non-negative. However, if  $t_1 = 10$ , then the weighting function  $\frac{t - t_1}{t_2 - t_1}$  will take negative values for  $t < 10$ . Similarly, if  $t_2 = 40$ , then the weighting function  $\frac{t_2 - t}{t_2 - t_1}$  will take negative values for  $t > 40$ . Hence, even if the  $\beta(t)$  is non-negative (or even strictly positive) across the whole life course, MVMR estimates can be negative. For instance, if  $t_1 = 10$  and  $t_2 = 50$ , then the asymptotic estimate at  $t_2$  would be negative if  $\beta(t)$  is large and positive for  $t < 10$ , and zero (or alternatively small and positive) for  $t \geq 10$ .

We note that the sum of the asymptotic estimates from MVMR is the total cumulative effect. This result holds more generally provided the genetic effects on the exposure are not too irregular (see [Supplementary Text S3](#), available as [Supplementary data](#) at *IJE* online for precise conditions).

We also consider simulated scenarios for a data-generating model in which the genetic effects are linear and the confounder is allowed to vary in time:

$$X(t) = \sum_{j=1}^{30} \alpha_j(t) G_j + U_0 + U(t) + \epsilon_X(t) \quad t \in [0, 50] \quad (7)$$

$$Y = \int_0^{50} \beta(t) X(t) dt + U_0 + U(50) + \epsilon_Y(50) \quad (8)$$

where  $U(t)$ ,  $\epsilon_X(t)$ ,  $\epsilon_Y(t)$  are independent Brownian motions

with variance equal to 1 at  $t = 50$ ,  $U_0 \sim \mathcal{N}(0, 1^2)$  and the instrumental effects on the exposure are:

$$\alpha_j(t) = a_j + b_j t \quad (9)$$

where  $\{a_j\}$  and  $\{b_j\}$  are independently simulated from the uniform distribution  $U(-0.1, 0.1)$  and  $U(-0.04, 0.04)$ , respectively. The trajectory for the exposure of one individual, along with the instrument and confounding effects, is shown in [Supplementary Figure S1](#) (available as [Supplementary data](#) at *IJE* online). We consider the true underlying effect  $\beta(t)$  of Scenario 2A (null effect until time 40 and positive effect thereafter) and run the MVMR analysis with four sets of time points; they are 10, 50, 20, 50, 40, 50 and 20, 40, denoted by Scenario 3A, 3B, 3C and 3D. For each case, we also calculate the expected MVMR estimates according to Equation (6). Finally, to investigate the sensitivity of our findings to the choice of genetic model, we conduct a further simulation in a scenario in which the effects of genetic variants on the exposure switch from off to on at different times (see [Supplementary Text S4](#), available as [Supplementary data](#) at *IJE* online, for full details).

### Illustrative example: BMI and SBP

To investigate how estimates may behave in a real data analysis, we performed a time-varying MVMR analysis in which the exposure is BMI and the outcome is SBP. Previous Mendelian randomization analyses have suggested that BMI has a positive causal effect on SBP,<sup>18</sup> although how this effect may vary over time has not been explored. We took data from UK Biobank, a prospective cohort study of around half a million people aged 40–69 years at baseline, recruited in 2006–2010 from across the UK.<sup>19</sup> We considered 366 089 unrelated individuals of European ancestries, who passed various quality-control filters as described previously.<sup>20</sup> We used 93 uncorrelated (pairwise  $r^2 < 0.01$ ) single-nucleotide polymorphisms as instrumental variables that have previously been shown to be associated with BMI at a genome-wide level of statistical significance.<sup>21</sup> This genome-wide association study did not include UK Biobank participants, thus avoiding bias due to winner's curse.<sup>22</sup>

We derived MVMR estimates using a two-sample 2SLS method, as we only have measurements of BMI for each individual at a single time point. In the first stage, we performed two separate regressions of BMI on the genetic variants using individuals recruited between ages 41–46 and 60–65 years. We used coefficients from these regressions to estimate genetically predicted BMI at ages 41–46 and 60–65 years for individuals aged  $>65$  years at recruitment. We then performed the second-stage regression of

SBP on genetically predicted BMI using individuals aged >65 years at recruitment. To investigate variability of results to the choice of time period, we repeated this analysis using different age ranges for the first BMI regression: 42–47, 43–48, 44–49 and so on up to 55–60 years. The second time period was always taken as 60–65 years, and the second-stage regression was always conducted in individuals aged >65 years at recruitment. Estimates represent the change in SBP in mmHg per 1-kg/m<sup>2</sup> increase in genetically predicted BMI. The diagram demonstrating the fitting procedure is shown in [Supplementary Figure S2](#) (available as [Supplementary data](#) at *IJE* online).

## Results

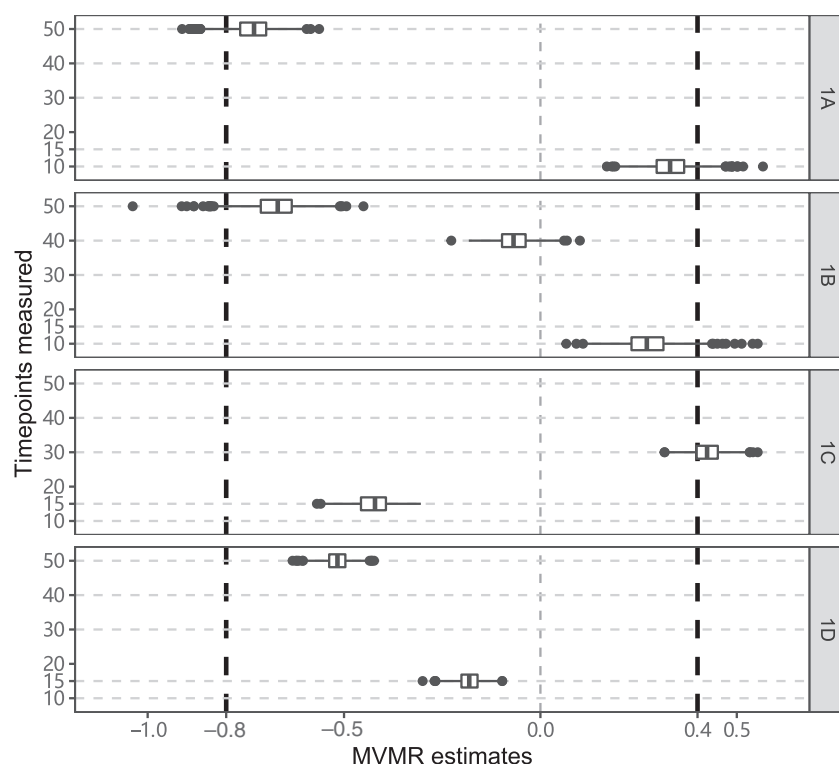
### Simulation study

For each scenario, we generated 1000 data sets on 10 000 participants. The average proportion of variance in the exposure explained by the genetic variants under each scenario is ~10%, corresponding to a univariable F-statistic of 36.9. The average conditional F-statistic values for MVMR in most scenarios are >10, a value conventionally regarded as a guide to diagnose weak instruments.<sup>4,11</sup> Detailed values for instrument strength values are shown in [Supplementary Table S2](#)

(available as [Supplementary data](#) at *IJE* online). We present results from MVMR obtained using the 2SLS method.

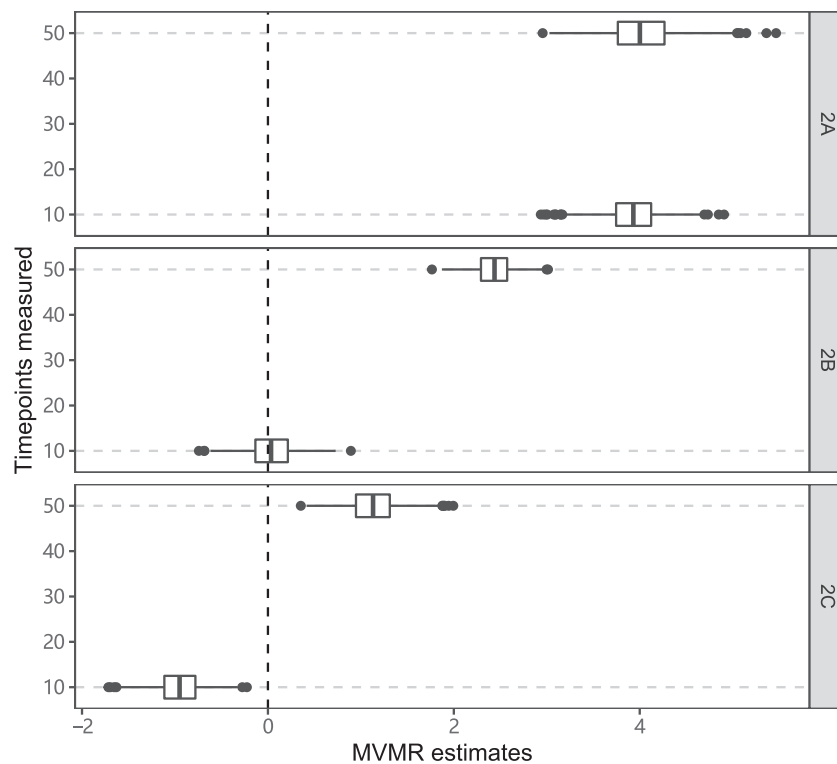
### Discrete effects of the exposure at specific time points

Results from the simulation study in which the outcome is a function of the exposure at times 10 and 50 are shown in [Figure 3](#). In Scenario 1A, we use the exposure measured at times 10 and 50 as risk factors. Median estimates across scenarios are close to the true causal effects: namely +0.4 at time 10 and –0.8 at time 50, though there is some bias towards the null due to weak instruments. In Scenario 1B, we use the exposure measured at times 10, 40 and 50 as risk factors. Again, median estimates across scenarios are close to the true causal effects, with the median estimate at time 40 around zero. However, in Scenario 1C (using the exposure at times 15 and 30) and in Scenario 1D (using the exposure at times 15 and 50), median estimates are substantially different to the true values. In Scenario 1C, the median estimate at time 15 is negative and at time 30 is positive; this is the opposite to the true situation, as the true effect is positive at the earlier time point and negative at the later time point. In Scenario 1D, the median estimate



**Figure 3** Simulation results when the outcome is affected discretely by the exposure at specific time points. Box plots of multivariable Mendelian randomization (MVMR) estimates with risk factors taken as the exposure measure at different measured time points—Scenario 1A: times 10 and 50; Scenario 1B: 10, 40, 50; Scenario 1C: 15, 30; Scenario 1D: 15, 50. Box indicates lower quartile, median and upper quartile; error bars represent the minimal and maximal data point falling in the 1.5 interquartile range distance from the lower/upper quartile; estimates outside this range are plotted separately. The true effects are  $\beta_1 = 0.4$  at time 10 and  $\beta_2 = -0.8$  at time 50 (black dashed lines)





**Figure 4** Simulation results when the outcome is affected continuously by the exposure across time. Box plots of multivariable Mendelian randomization (MVMR) estimates with risk factors taken as the exposure measure at different measured time points: Scenario 2A: null effect until time 40, positive effect thereafter; Scenario 2B: positive effect until time 20, null effect thereafter; Scenario 2C: constant positive effect. Box indicates lower quartile, median and upper quartile; error bars represent the minimal and maximal data point falling in the 1.5 interquartile range distance from the lower/upper quartile; estimates outside this range are plotted separately

at time 50 is correctly negative, but the median estimate at time 15 is also negative, whereas the early-age causal effect is positive. We repeated the simulation study with stronger instruments (average proportion of variance in the exposure explained was around 30%); results shown in [Supplementary Figure S3](#) (available as [Supplementary data](#) at *IJE* online) support our claim that bias in Scenarios 1A and 1B is due to weak instruments.

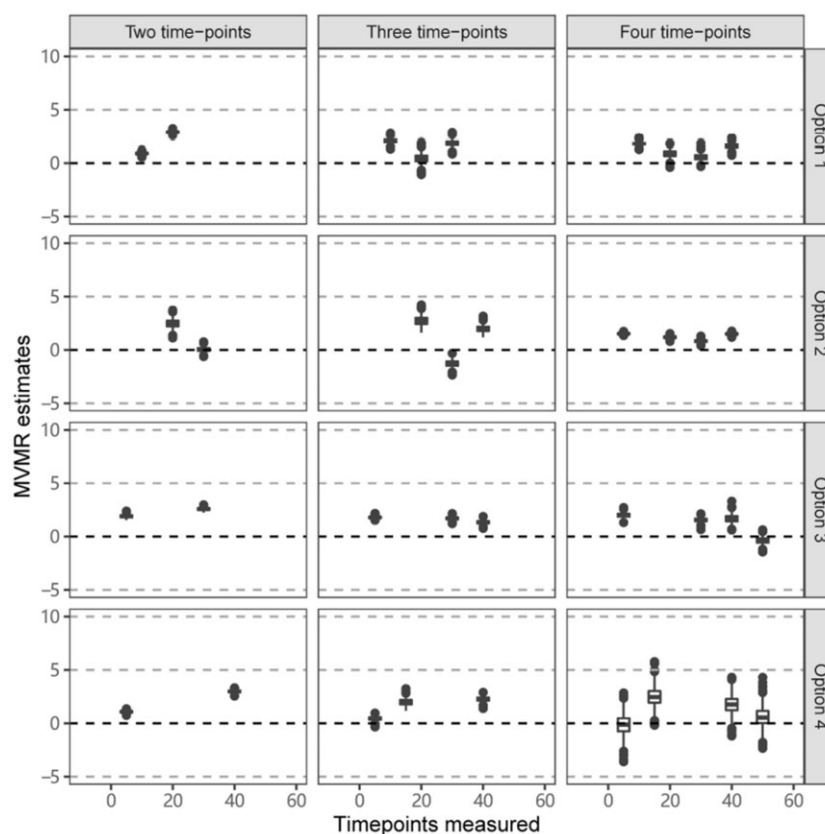
### Continuous effects of the exposure across time

Results from the simulation study in which the outcome is a continuous function of the exposure are shown in [Figure 4](#). In Scenario 2A, median estimates at both time points are positive despite the true effect being null in early life and positive in later life only. In Scenario 2B, the median estimate at the earlier time point is close to zero and the median estimate at the later time point is positive, whereas in truth the causal effect of the exposure is positive in early life and null in later life. In Scenario 2C, the median estimate at the earlier time point is negative and the median estimate at the later time point is positive, despite the true causal effect of the exposure being constant and positive throughout.

[Figure 5](#) provides results from Scenario 2C (constant positive effect) for a range of different choices of time points, indicated as Option 1, Option 2 and so on. We also vary the number of time points considered. Median estimates varied substantially depending on the choice of time points, with no evident pattern of results. Although median estimates were mostly positive, they were sometimes close to zero and occasionally clearly negative. Moreover, it is not clear that estimates improved when considering measures of the exposure at increased numbers of time points, as some median estimates were either close to zero or negative even when the exposure was measured at four time points.

### Linear time-varying genetic effects

Results from the simulation study in which the outcome is a continuous function of the exposure and the genetic effects on the exposure vary linearly with time are shown in [Figure 6](#). In Scenarios 3A and 3B, the median estimates for the first measured time point are positive, whereas in truth the causal effect of the exposure is null at these time points. When the second time point is changed to 40, as in Scenario 3D, the median estimates for the first measured



**Figure 5** Simulation results when the outcome is affected continuously by the exposure across time in Scenario 2C (constant positive effect). Box plots of multivariable Mendelian randomization (MVMR) estimates with risk factors taken as the exposure measure at different measured time points when varying the location and number of time points. Box indicates lower quartile, median and upper quartile; error bars represent the minimal and maximal data point falling in the 1.5 interquartile range distance from the lower/upper quartile; estimates outside this range are plotted separately.

time point are negative, despite the true causal effect function being non-negative at all time points. In all scenarios, the median estimates are close to the expected values given in Equation (6). This indicates that the problem in time-varying MVMR is not bias, but rather that the quantities estimated do not have a meaningful interpretation as the causal effect of an exposure restricted to a particular time period.

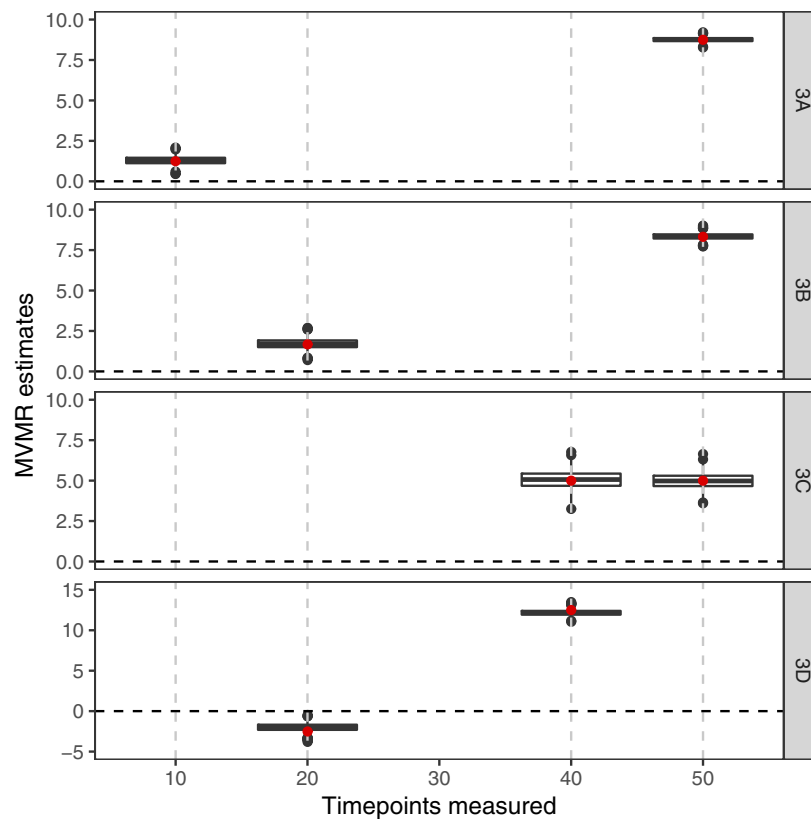
Finally, results from the simulation in which the effects of genetic variants on the exposure switch from off to on at different times are presented in Supplementary Figures S5 and S6 (available as Supplementary data at *IJE* online). Results are very similar to those in Figure 6, indicating that findings are not sensitive to the choice of model relating the genetic variants to the exposure.

### Illustrative example: BMI and SBP

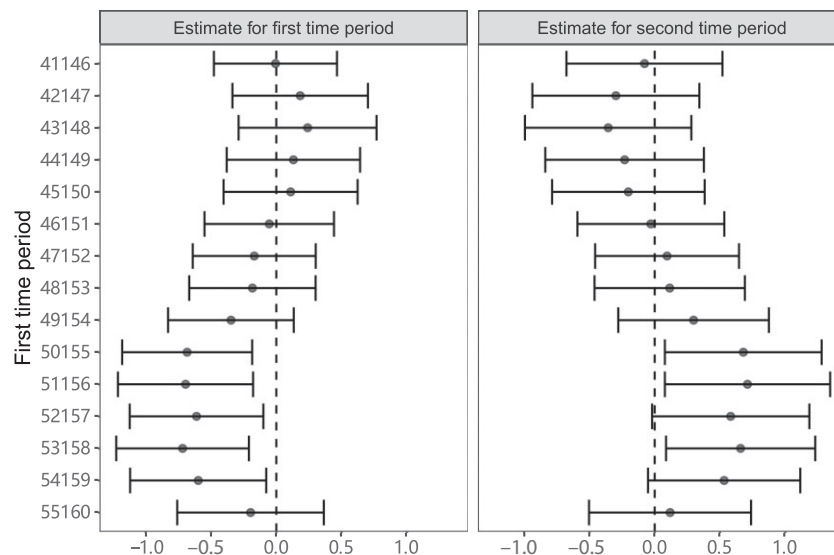
Results from the illustrative example for the effect of BMI on SBP are shown in Figure 7. Conditional F-statistics were close to 1 in all cases, indicating little variability in

the genetic effect over time (Supplementary Table S3, available as Supplementary data at *IJE* online). This means that there is little information in the data to obtain precise multivariable estimates at the different time points. Scatter plots of the genetic associations with BMI at different time periods are shown in Supplementary Figure S4 (available as Supplementary data at *IJE* online); whereas most variants are similarly associated with BMI across time periods, some variants are more strongly associated during the earlier or later time period, enabling the MVMR analysis to be performed. MVMR estimates for both time periods are highly variable about the age range over which the risk factor associations were estimated. The range of variation in Mendelian randomization estimates was greater than expected based on the standard errors of estimates. In some cases, the 95% CI for the estimate in the first time period was negative and excluded the null and the 95% CI for the estimate in the second time period was positive and excluded the null. But there were also positive point estimates for the first time period and negative point estimates for the second time period. In summary, both estimates





**Figure 6** Simulation results when the outcome is affected continuously by the exposure across time and the genetic effects on the exposure vary linearly in time. Box plots of multivariable Mendelian randomization (MVMR) estimates with risk factors taken as the exposure measure at different measured time points: 10, 50, 20, 50, 40, 50 and 20, 40, denoted by Scenario 3A, 3B, 3C and 3D. The underlying continuous effect is as in Scenario 2A: null effect until time 40, positive effect thereafter. The red points represent the expected estimate according to Equation (6). Box indicates lower quartile, median and upper quartile; error bars represent the minimal and maximal data point falling in the 1.5 interquartile range distance from the lower/upper quartile; estimates outside this range are plotted separately



**Figure 7** Multivariable Mendelian randomization estimates (95% CIs) of the effect of body mass index (BMI) on systolic blood pressure for different time periods. The first risk factor is BMI over the first time period as indicated. The second risk factor is BMI over the time period from age 60 to 65 years. Estimates for systolic blood pressure are performed in individuals aged >65 years at the time of recruitment

and inferences based on those estimates were strongly dependent on the choice of the first time period—a choice that is likely to be arbitrary in practice.

## Discussion

As the well-known aphorism goes: all models are incorrect, but some models are useful.<sup>23</sup> In univariable Mendelian randomization, strict parametric assumptions are required for Mendelian randomization estimates to be interpreted as either an ATE or a LATE.<sup>12</sup> However, even if these assumptions are not satisfied, the univariable Mendelian randomization estimate still has an interpretation as a test statistic for a relevant causal hypothesis.<sup>24</sup> Indeed, it has been argued that the numerical value of a Mendelian randomization estimate rarely represents the estimate of a policy-relevant parameter,<sup>25</sup> as (amongst other reasons) it represents the impact of a lifelong change in the distribution of the exposure, whereas interventions on exposures in clinical practice are more limited in time.<sup>26,27</sup> Hence the primary value of a Mendelian randomization investigation is to provide evidence supporting (or questioning) a causal hypothesis, rather than providing a causal estimate.<sup>28</sup>

However, estimates from MVMR for a time-varying exposure do not seem to have a similar interpretation as a test statistic for a relevant causal hypothesis relating to the presence of a causal effect over a specific time period. When the exposure affects the outcome at a limited number of discrete time points and the risk factors in the MVMR analysis are the values of the exposure at these time points, causal effects at these time points can be unbiasedly estimated. But if these time points are not correctly identified, estimates obtained by the ill-specified model are incorrect in a way that is misleading to any inferences being drawn from their magnitude, about either the presence or the direction of a causal effect. Similarly, if the effect of the exposure on the outcome is not discrete, but rather continuous, then the values estimated by MVMR do not have a naturally intuitive interpretation. For the MVMR model to be correctly specified, the model in [Figure 1](#) must represent the true data-generating model, and not merely be a simplification of the model. It is implausible for the true effect of an exposure on the outcome to be discrete at the precise time points at which measurements of the exposure are available; hence, we expect potentially misleading estimates to be ubiquitous for time-varying Mendelian randomization investigations in practice.

We have provided a simulated example in which the effect of the exposure on the outcome is positive in early life only and negative in later life, but estimates from MVMR would suggest the opposite. We have also provided a

simulated example in which the effect of the exposure on the outcome occurs in early life only and not in later life, but estimates from MVMR would suggest the opposite. Similarly, we have provided a simulated example in which the effect of the exposure on the outcome occurs in later life only and not in early life, but estimates from MVMR analysis are positive throughout. We have provided a simulated example in which estimates from MVMR at different time points are not necessarily positive even though the effect of the exposure is uniformly constant and positive. Finally, we have considered a simple case in which the genetic effects on the exposure vary linearly, and have shown algebraically that values estimated by MVMR do not represent intuitively interpretable causal effects. We have also provided an applied example with real data, and shown that estimates are sensitive about the time period over which genetic associations with the exposure were estimated. The choice of time period influenced the conclusions drawn from the analysis: whether there is a negative effect of BMI on SBP during the first time period or not, and whether there is a positive effect of BMI on SBP during the second time period or not.

The poor performance of MVMR in these examples can be attributed to model misspecification. In general, when a statistical model is misspecified, estimates suffer from bias that is unpredictable in both magnitude and direction.<sup>29</sup> The estimates of MVMR represent weighted cumulative effects, whereas the weighting function is affected by the measured time points and the time-varying instrumental effects, which makes results sensitive to the time points chosen and unpredictable as the time-varying instrument effects are unknown. Another explanation is violation of the exclusion restriction assumption. The instrumental variable assumptions state that the totality of the effect of the genetic variants on the outcome is mediated via the exposure, such that if the exposure were fixed and the genetic variants varied, the outcome would remain the same.<sup>30</sup> If the effect of the exposure acts continuously over time, it would not be sufficient to fix its value at a few specific time points, but it would be necessary to fix the trajectory of the exposure over time. Similarly, complete mediation of the genetic effect on the outcome via the exposure would only be observed when considering the distribution of the exposure over time. As time is continuous rather than discrete, it is unlikely that any model containing values of the exposure at specific discrete time points could be correctly specified. One potential extension of this work is to consider models that fit the outcome as a continuous function of the exposure, which requires the exposure or the instrument effects over the whole life course to be modelled.

One qualitative limitation of our cautionary remarks is that we have considered scenarios in which the risk factors represent the same exposure measured at different time points. In the example of Richardson *et al.*,<sup>7</sup> it is arguable that early-life BMI and later-life BMI represent biologically distinct exposures, whose effect on the outcome is restricted to a limited time period. For example, early-life BMI may only affect the outcome for the growth and development period until full adult size is achieved. Hence MVMR could be a legitimate tool for obtaining meaningful inferences in this case, and the estimates represent the effect of the corresponding distinct exposure during the time period. Whereas we agree that such a situation is more amenable to the use of MVMR, we would still be cautious about over-interpretation of results from time-varying Mendelian randomization investigations, and would recommend that investigators assess the robustness of findings carefully (e.g. by assessing the consistency of results with different choices of genetic variants and/or different time points for the measurement of the exposure). Another limitation is that the simulation scenarios considered may be unrealistic, although similar fragility in estimates was observed in the applied analysis. A further practical issue, as observed in the applied example, is that it is difficult to find genetic variants for which the effect on the exposure is more or less strong at different time periods, and hence estimates from MVMR may be highly variable and imprecise.

In conclusion, MVMR analyses to investigate time-varying causal effects rely on parametric assumptions that are unlikely to be satisfied in practice, and provide estimates that can be misleading if the model is incorrect. We therefore strongly discourage quantitative conclusions to be drawn from these analyses, and would advise caution about the qualitative interpretation of such findings as a guide of the direction or even existence of a causal effect during a particular phase of life.

## Ethics approval

The UK Biobank study has approval from the North West Multicentre Research Ethics Committee (11/NW/0382).

## Data availability

Individual-level data from UK Biobank cannot be shared publicly for ethical/privacy reasons. The data will be shared on reasonable request to the corresponding author, with the permission of UK Biobank. The research has been conducted using the UK Biobank resource under Application Number 7439.

## Supplementary data

Supplementary data are available at *IJE* online.

## Author contributions

H.T. designed and conducted the simulation and applied study. S.B. directed the study's implementation and helped to interpret the findings. All authors reviewed and edited the manuscript.

## Funding

This research was supported by the United Kingdom Research and Innovation Medical Research Council (MC\_UU\_00002/7) and the National Institute for Health Research Cambridge Biomedical Research Centre (BRC-1215–20014). S.B. is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (204623/Z/16/Z). The views expressed are those of the authors and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care.

## Acknowledgements

We thank Christopher Jackson and Jessica Barrett (MRC Biostatistics Unit) for helpful discussions in the development of this work. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

## Conflict of interest

None declared.

## References

1. Burgess S, Thompson S. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* 2015;181:251–60.
2. Burgess S, O'Donnell CJ, Gill D. Expressing results from a Mendelian randomization analysis: separating results from inferences. *JAMA Cardiol* 2021;6:7–8.
3. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res* 2007;16:309–30.
4. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single sample and two-sample summary data settings. *Int J Epidemiol* 2019;48:713–27.
5. Carter AR, Sanderson E, Hammerton G *et al.* Mendelian randomisation for mediation analysis: current methods and challenges for implementation. *Eur J Epidemiol* 2021;36:465–78.
6. Burgess S, Thompson D, Rees J, Day F, Perry J, Ong K. Dissecting causal pathways using Mendelian randomization with summarized genetic data: application to age at menarche and risk of breast cancer. *Genetics* 2017;207:481–87.
7. Richardson TG, Sanderson E, Elsworth B, Tilling K, Davey Smith G. Use of genetic variation to separate the effects of early and later life adiposity on disease risk: Mendelian randomisation study. *Br Med J* 2020;369:m1203.

8. Gill D, Georgakis M, Zuber V *et al.* Genetically predicted midlife blood pressure and coronary artery disease risk: Mendelian randomization analysis. *J Am Heart Assoc* 2020;**9**:e016773.
9. Sanderson E, Richardson TG, Morris TT, Tilling K, Davey Smith G. Estimation of causal effects of a time-varying exposure at multiple time points through multivariable Mendelian randomization. *PLoS Genet* 2022;**18**:e1010290.
10. Shi J, Swanson SA, Kraft P, Rosner B, De Vivo I, Hernán MA. Mendelian randomization with repeated measures of a time-varying exposure: an application of structural mean models. *Epidemiology* 2022;**33**:84–94.
11. Sanderson E, Windmeijer F. A weak instrument F-test in linear IV models with multiple endogenous variables. *J Econ* 2016;**190**:212–21.
12. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology* 2006;**17**:360–72.
13. Swanson SA, Hernán MA. Commentary: how to report instrumental variable analyses (suggestions welcome). *Epidemiology* 2013;**24**:370–74.
14. Swanson SA, Labrecque J, Hernán MA. Causal null hypotheses of sustained treatment strategies: what can be tested with an instrumental variable? *Eur J Epidemiol* 2018;**33**:723–28.
15. Burgess S, Thompson SG. *Mendelian Randomization: Methods for Causal Inference Using Genetic Variants*. Chapter 9, Extensions to Mendelian randomization, 2nd ed. New York: Chapman & Hall/CRC, 2021, pp. 149–166.
16. Wooldridge JM. Introductory econometrics: a modern approach. Chapter 15, Instrumental Variables Estimation and Two Stage Least Squares, 4th ed. South-Western, Nashville, TN: South Western College Publications, 2009, pp. 506–543.
17. Burgess S, Dudbridge F, Thompson SG. Re: 'Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects'. *Am J Epidemiol* 2015;**181**:290–91.
18. Holmes MV, Lange LA, Palmer T *et al.* Causal effects of body mass index on cardiometabolic traits and events: a Mendelian randomization analysis. *Am J Hum Genet* 2014;**94**:198–208.
19. Sudlow C, Gallacher J, Allen N *et al.* UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;**12**:e1001779.
20. Astle WJ, Elding H, Jiang T *et al.* The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* 2016;**167**:1415–29.e19.
21. Locke AE, Kahali B, Berndt SI *et al.*; International Endogene Consortium. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;**518**:197–206.
22. Taylor AE, Davies NM, Ware JJ, VanderWeele T, Smith GD, Munafò MR. Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. *Econ Hum Biol* 2014;**13**:99–106.
23. Box GE. Science and statistics. *J Am Stat Assoc* 1976;**71**:791–99.
24. Labrecque JA, Swanson SA. Interpretation and potential biases of Mendelian randomization estimates with time-varying exposures. *Am J Epidemiol* 2019;**188**:231–38.
25. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in Mendelian randomization. *Epidemiology* 2014;**25**:427–35.
26. Burgess S, Butterworth A, Malarstig A, Thompson S. Use of Mendelian randomisation to assess potential benefit of clinical intervention. *BMJ* 2012;**345**:e7325.
27. Ference BA. How to use Mendelian randomization to anticipate the results of randomized trials. *Eur Heart J* 2018;**39**:360–62.
28. Burgess S, Davey Smith G, Davies NM *et al.* Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res* 2019;**4**:186.
29. Vansteelandt S, Bekaert M, Claeskens G. On model selection and model misspecification in causal inference. *Stat Methods Med Res* 2012;**21**:7–30.
30. Clarke PS, Windmeijer F. Instrumental variable estimators for binary outcomes. *J Am Stat Assoc* 2012;**107**:1638–52.